Sparse Bayesian pMRI Reconstruction With Complex Bernoulli-Laplace Mixture Priors

Siwar Chaabene[1,2], Lotfi Chaari[1,2] and Abdelaziz Kallel[1]

[1] Digital Research Centre of Sfax, Tunisia.

[2] MIRACL laboratory, University of Sfax, Tunisia

***Abstract*—This paper presents a sparse Bayesian regularization technique for image restoration in parallel magnetic resonance imaging (pMRI). This technique is based on a hierarchical Bayesian model that solves the inverse problem of pMRI reconstruction by promoting sparsity using a Bernoulli-Laplace mixture prior. A Markov Chain Monte Carlo (MCMC) sampling technique is used to numerically approximate the target posterior. Our model allows handling complex-valued data. Promising results obtained on synthetic data demonstrate the performance of the proposed sparse Bayesian restoration model to provide accurate estimation of the target images.**

***Index Terms*—Sparse Bayesian model, MCMC, parallel MRI restoration**

I. Introduction

Parallel imaging [1] with several receiver coils having different spatial sensitivity profiles has been the major innovation in magnetic resonance imaging (MRI) since the early 1990s. It allows faster acquisition of MRI images and improves its spatial and temporal resolutions. Thus, the principle of reconstruction in parallel imaging consists in combining images of several receiver coils in order to reconstruct a global image. Consequently, SENSitivity Encoding (SENSE) [2] is the most robust commonly used reconstruction technique in the clinical routine compared to the other existing techniques [3].

In this study, we focus on the SENSE reconstruction technique operating in the spatial domain. Therefore, parallel MRI reconstruction/restoration based on the SENSE technique is considered as an ill-posed inverse problem [4]. In practice, the reconstructed images by SENSE are often tainted by severe artifacts caused by the lack of precision in the sensitivity maps estimation, the presence high level of the observation noise and especially the use of high reduction factors.

On one hand, the resolution of this inverse problem regularized SENSE methods has made significant progress during the last decade [4]–[7]. On the other hand, the use of $\ell_0 + \ell_1$ regularization for sparse signal and image recovery has generated research interest in order to resolve the ill-posed inverse problem in diverse areas such as biomedical imaging reconstruction. Therefore, the inject priors based on mixtures of Bernoulli and Laplace distributions in the observation model allow to use $\ell_0 + \ell_1$ norms regularization resulting in a valid estimation for the sparsity of the desired image as illustrated in the recent works [8]–[10]. The Bernoulli-Laplace (BL) based models developed in a Bayesian framework allow the regularization parameters/hyperparameters to be automatically estimated based on the observe data.

In this paper, we develop a sparse Bayesian regularization technique for the complex-valued pMRI reconstruction based on $\ell_0 + \ell_1$ norm priors for the estimation of the target image with sensitivity errors during the reconstruction process.

The rest of this paper is divided into five sections. Section II introduces the problem formulation of the parallel MRI reconstruction. Section III details the proposed Bayesian sparse technique for pMRI restoration. Section IV shows the adopted inference scheme. Section V dresses an experimental validation on a complex-valued synthetic dataset. Finally, Section VI presents conclusions and some perspectives.

II. Problem Formulation

The linear observation model of parallel MRI [5] in the image domain at each spatial position $\boldsymbol{x}$ based on the SENSE method is modeled as

$$\boldsymbol{d}(\boldsymbol{x}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{\rho}(\boldsymbol{x}) + \boldsymbol{n}(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{d}$ represents the observation signal, $\boldsymbol{\rho}$ is the target image to be estimated, $\boldsymbol{S}$ represent the sensitivity maps matrix and $\boldsymbol{n}$ is the additive Gaussian observation noise.

For this case of inverse problem, the sensitivity maps operator $\boldsymbol{S}$ is ill-conditioned. However, the pMRI reconstruction is considered an ill-posed inverse problem. The main objective of our work is to accurately estimate the desired image $\boldsymbol{\rho}$ from the observation $\boldsymbol{d}$ with the presence of errors in the sensitivity maps.

The following section presents the sparse Bayesian model applied for regularized SENSE reconstruction while taking into account the sensitivity maps errors. Note that, the pMRI data are complex-values, and $\boldsymbol{d}$, $\boldsymbol{\rho}$ and $\boldsymbol{S}$ are assumed to be realizations of random variables.

III. Hierarchical Bayesian model

A. Likelihood

The acquisition noise in pMRI is assumed to be complex-valued, additive and Gaussian. From this assumption, and based on the observation model (1), the likelihood function can be written as follows:

$$f(\boldsymbol{d}|\boldsymbol{\rho}, \sigma_n^2) = \frac{\prod_{\boldsymbol{x}} \exp\left(-\|\boldsymbol{d}(\boldsymbol{x}) - \boldsymbol{S}(\boldsymbol{x})\boldsymbol{\rho}(\boldsymbol{x})\|^2_{\Psi^{-1}}\right)}{(2\pi)^{M/2}|\Psi|^{1/2}}, \tag{2}$$

where M represent the number of pixels and Ψ^{-1} is the noise covariance matrix. For the sake of simplicity, we can assume that $\Psi = \sigma_n^2 I$, where I is the identity matrix.

B. Prior distributions

We detail here the prior distributions retained for the unknown parameter vector $\boldsymbol{\theta} = \left\{\sigma_n^2, \boldsymbol{\rho}\right\}$ to be estimated.

1) Prior distribution for σ_n^2:

We use a conjugate non-informative prior distribution to guarantee the positivity for the noise variance. Specifically, an inverse-Gamma distribution with hyperparameters α and β is used

$$f(\sigma_n^2|\alpha,\beta) = \mathcal{IG}(\sigma_n^2|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\sigma_n^2\right)^{-1-\alpha}\ \exp\left(-\frac{\beta}{\sigma_n^2}\right), \tag{3}$$

where $\Gamma(.)$ is the gamma function. In order to guarantee a non-informative prior, the hyperparameters α and β are fixed by 10^{-3}.

2) Prior for $\boldsymbol{\rho}$:

In order to promote the sparsity of the target image in the original domain, we propose to use a Bernoulli-Laplace (BL) mixture distribution. Since the processed data is complex-valued, a BL model for the real and imaginary parts is used separately. Assuming that every coefficient $\rho_i (i = 1, ..., N)$ can be expressed as $\rho_i = a_i + jb_i$ (N denotes the number of pixels in the target image $\boldsymbol{\rho}$ and $j^2 = -1$) where a_i and b_i denote the real and imaginary parts of ρ_i, respectively, and assuming that the real and imaginary parts are independent, the used prior can be expressed as:

$$\begin{aligned} f(\rho_i|\omega,\lambda) &= f(a_i|\omega,\lambda) f(b_i|\omega,\lambda) \\ &= \left[(1-\omega)\,\delta(a_i) + \omega\frac{1}{2\lambda}\exp\left(-\frac{|a_i|}{\lambda}\right)\right] \\ &\times \left[(1-\omega)\,\delta(b_i) + \omega\frac{1}{2\lambda}\exp\left(-\frac{|b_i|}{\lambda}\right)\right], \end{aligned} \tag{4}$$

where $\delta(.)$ denotes Dirac delta function. As regards ω, it denotes a weight belonging to $[0, 1]$ that indicates the rate of non-zero coefficients. The hyperparameter λ assesses the sparsity level of the non-zero coefficients of both the real and imaginary parts. It is worth noticing that different hyperparameters could also be considered for each part.

In addition, we assume that the random variables ρ_i are independent and squeeze out the BL prior of the target image $\boldsymbol{\rho}$ as following :

$$f(\boldsymbol{\rho}|\omega,\lambda) = \prod_{i=1}^{N} f(\rho_i|\omega,\lambda). \tag{5}$$

C. Hyperprior distributions

Our hierarchical model is build upon two layers. The second level of hierarchy involves the hyperprior distributions for the unknown hyperparameter vector denoted by $\boldsymbol{\Phi} = \{\omega, \lambda\}$. This subsection defines the hyperprior choice for these two hyperparameters.

1) Hyperprior for ω:

For this hyperparameter, a uniform distribution on $[0, 1]$ is adopted:

$$f(\omega) = \mathcal{U}_{[0,1]}(\omega)\,. \tag{6}$$

Such a distribution helps keeping a non-informative prior. However, it is worth noticing that a more informative version could be considered if further informations about the non-zero coefficients rate in the target signal.

2) Hyperprior for λ:

We use a conjugate Inverse-gamma distribution $\mathcal{IG}(\lambda|\kappa,\vartheta)$ for the hyperparameter λ, where κ and ϑ has been set to 10^{-1} in order to keep a non-informative prior:

$$\mathcal{IG}(\lambda|\kappa,\vartheta) = \frac{\vartheta^\kappa}{\Gamma(\kappa)}\lambda^{-1-\kappa}\ \exp\left(-\frac{\vartheta}{\lambda}\right). \tag{7}$$

IV. Bayesian Inference Scheme

Based on the hierarchical Bayesian model detailed in Section III, we use a maximum a posterior (MAP) strategy to derive estimators for the model parameters and hyperparameters. According to the Bayes' theorem, the joint posterior distribution of $\{\boldsymbol{\theta}, \boldsymbol{\Phi}\}$ is proportional to the combination of the likelihood and the priors distributions, and can be written as:

$$\begin{aligned} f(\boldsymbol{\theta},\boldsymbol{\Phi}|\boldsymbol{d}) &\propto f(\boldsymbol{d}|\boldsymbol{\theta})\, f(\boldsymbol{\theta}|\boldsymbol{\Phi})\, f(\boldsymbol{\Phi}|\alpha,\beta,\kappa,\vartheta) \\ &\propto f(\boldsymbol{d}|\boldsymbol{\rho},\sigma_n^2) f(\boldsymbol{\rho}|\omega,\lambda) f(\sigma_n^2|\alpha,\beta) \\ &\times f(\omega) f(\lambda|\kappa,\vartheta). \end{aligned} \tag{8}$$

Using the distributions adopted in the previous section, the above posterior has the following form:

$$\begin{aligned} f(\boldsymbol{\theta},\boldsymbol{\Phi}|\boldsymbol{d}) \propto & \frac{\prod_{\boldsymbol{x}} \exp\left(-\|\boldsymbol{d}(\boldsymbol{x}) - \boldsymbol{S}(\boldsymbol{x})\boldsymbol{\rho}(\boldsymbol{x})\|^2_{\Psi^{-1}}\right)}{(2\pi)^{M/2}|\Psi|^{1/2}} \\ & \times \prod_{i=1}^{N}\left[\left((1-\omega)\delta(a_i) + \omega\frac{1}{2\lambda}\exp(-\frac{|a_i|}{\lambda})\right)\right. \\ & \left.\times\left((1-\omega)\delta(b_i) + \omega\frac{1}{2\lambda}\exp(-\frac{|b_i|}{\lambda})\right)\right] \\ & \times \mathcal{U}(\omega) \times \frac{\vartheta^\kappa}{\Gamma(\kappa)}(\lambda)^{-1-\kappa}\exp\left(-\frac{\vartheta}{\lambda}\right). \end{aligned} \tag{9}$$

The high complexity of this joint posterior do not allow deriving closed-form expressions of the target estimators. Hence, we use an MCMC technique to numerically approximate the target posterior. Specifically, we use a Gibbs sampler [11] which proceeds by sequential sampling according to the conditional distributions detailed below.

The proposed algorithm is executed repeatedly until establishing the convergence in order to provide a correct estimation for σ_n^2, ω, λ and $\boldsymbol{\rho}$. It worth to note, in our experience, we need 60 runs with 30 burn-in runs of our algorithm to ensure convergence.

Algorithm 1: Gibbs Sampler algorithm for Sparse pMRI Reconstruction.

Initialize $\boldsymbol{\rho}^{(\mathbf{0})}$;
repeat
  Sample σ_n^2 from $f(\sigma_n^2|\boldsymbol{d},\boldsymbol{\rho},\alpha,\beta)$.
  Sample ω from $f(\omega|\boldsymbol{\rho})$.
  Sample λ from $f(\lambda|\boldsymbol{\rho},\kappa,\vartheta)$.
  for $i = 1 \dots N$ **do**
    Sample a_i from $f(a_i|\boldsymbol{d},\tilde{\boldsymbol{\rho}}_i,\sigma_n^2,\omega,\lambda)$.
    Sample b_i from $f(b_i|\boldsymbol{d},\tilde{\boldsymbol{\rho}}_i,\sigma_n^2,\omega,\lambda)$.
  end
until *convergence*;

A. Sampling from $f(\sigma_n^2|\boldsymbol{d},\boldsymbol{\rho},\alpha,\beta)$

The conditional posterior distribution of σ_n^2 gives an Inverse-gamma distribution defined as:

$$\sigma_n^2|\boldsymbol{d},\boldsymbol{\rho},\alpha,\beta \sim IG\left(\alpha+\frac{M}{2},\beta+\frac{\|\boldsymbol{d}-\boldsymbol{S}\boldsymbol{\rho}\|_2^2}{2}\right), \quad (10)$$

where $\|.\|$ denotes the Euclidean norm. This conditional posterior distribution is easy to sample.

B. Sampling λ from $f(\lambda|\boldsymbol{\rho},\kappa,\vartheta)$

The conditional posterior distribution of λ is the following:

$$\lambda|\boldsymbol{\rho},\kappa,\vartheta \sim IG\left(\kappa+\|\boldsymbol{\rho}\|_0,\vartheta+\|\boldsymbol{\rho}\|_1\right). \quad (11)$$

where $\|.\|_0$ and $\|.\|_1$ refer to the ℓ_0 pseudo-norm and the ℓ_1 norm, respectively. This conditional posterior distributions is easy to sample.

C. Sampling from $f(\omega|\boldsymbol{\rho})$

The calculation of the posterior conditional distribution of ω gives the following beta distribution :

$$\omega|\boldsymbol{\rho} \sim B\left(1+\|\boldsymbol{\rho}\|_0,1+M-\|\boldsymbol{\rho}\|_0\right). \quad (12)$$

This conditional posterior distribution is easy to sample.

D. Sampling from $f(\boldsymbol{\rho}|\boldsymbol{d},\sigma_n^2,\omega,\lambda)$

Note that, the real and imaginary parts of ρ_i are sampled separately. The associated conditional posteriors are given respectively by:

$$\begin{aligned} f(a_i|\boldsymbol{d},\tilde{\boldsymbol{\rho}}_i,\sigma_n^2,\omega,\lambda) &= \omega_{1,i}^a\delta(a_i)+\omega_{2,i}^a\mathcal{N}^+\left(\mu_{a,i}^+,\sigma_i^2\right) \\ &\quad +\omega_{3,i}^a\mathcal{N}^-\left(\mu_{a,i}^-,\sigma_i^2\right), \end{aligned} \quad (13)$$

and

$$\begin{aligned} f(b_i|\boldsymbol{d},\tilde{\boldsymbol{\rho}}_i,\sigma_n^2,\omega,\lambda) &= \omega_{1,i}^b\delta(b_i)+\omega_{2,i}^b\mathcal{N}^+\left(\mu_{b,i}^+,\sigma_i^2\right) \\ &\quad +\omega_{3,i}^b\mathcal{N}^-\left(\mu_{b,i}^-,\sigma_i^2\right), \end{aligned} \quad (14)$$

where $\mathcal{N}^+$ (resp. $\mathcal{N}^-$) denote the truncated Gaussian distribution on $\mathbb{R}^+$ (resp. $\mathbb{R}^-$).
The target image $\boldsymbol{\rho}$ decompose onto the orthonormal basis $B=\{e_1,...,e_N\}$ such that $\boldsymbol{\rho}=\tilde{\boldsymbol{\rho}}_i+\rho_i\boldsymbol{e}_i$ where $\tilde{\boldsymbol{\rho}}_i$ is obtained by setting the i^{th} element of $\boldsymbol{\rho}$ to 0 and $\boldsymbol{v}_i=\boldsymbol{d}-\boldsymbol{S}\tilde{\boldsymbol{\rho}}_i$ and $\boldsymbol{s}_i=\boldsymbol{S}\boldsymbol{e}_i$.

The weights $\left(\omega_{l,i}^a\right)_{1\le l\le 3}$ used in (13) are determined as

$$\omega_{l,i}^a=\frac{u_{l,i}^a}{\sum_{l=1}^{3}u_{l,i}^a}, \quad (15)$$

where

$$\begin{aligned} u_{1,i}^a &= 1-\omega, \\ u_{2,i}^a &= \tfrac{\omega}{4\lambda^2}\exp\left(\tfrac{\left(\mu_{a,i}^+\right)^2}{2\sigma_i^2}\right)\sqrt{2\pi\sigma_i^2}C\left(\mu_{a,i}^+,\sigma_i^2\right), \\ u_{3,i}^a &= \tfrac{\omega}{4\lambda^2}\exp\left(\tfrac{\left(\mu_{a,i}^-\right)^2}{2\sigma_i^2}\right)\sqrt{2\pi\sigma_i^2}C\left(-\mu_{a,i}^-,\sigma_i^2\right), \end{aligned}$$

and

$$\begin{aligned} \sigma_i^2 &= \tfrac{\sigma_n^2}{\|\boldsymbol{s}_i\|_2^2}, \\ \mu_{a,i}^+ &= \sigma_i^2\left(\tfrac{Real\left(\boldsymbol{v}_i^T\boldsymbol{s}_i\right)}{\sigma_n^2}-\tfrac{1}{\lambda}\right), \\ \mu_{a,i}^- &= \sigma_i^2\left(\tfrac{Real\left(\boldsymbol{v}_i^T\boldsymbol{s}_i\right)}{\sigma_n^2}+\tfrac{1}{\lambda}\right), \\ C\left(\mu,\sigma^2\right) &= \sqrt{\tfrac{\sigma^2\pi}{2}}\left(1+erf\left(\tfrac{\mu}{\sqrt{2\sigma^2}}\right)\right). \end{aligned}$$

Similar expressions are obtained for $\left(\omega_{l,i}^b\right)_{1\le l\le 3}$.

V. Experimental Validation

To demonstrate the efficiency of the proposed sparse Bayesian reconstruction model, we apply our algorithm to synthetic data of two different brain slices of size 256×256. We simulate poor acquisition conditions by using $L=8$ receiver coils and a reduction factor $R=4$. Moreover, we simulate errors in the estimation of the coil sensitivity maps by adding a Gaussian white noise with a variance equal to the value 0.001. After simulating the acquisition process, a complex Gaussian noise with variance $\sigma_n^2=4$ is added to each simulated MRI data.
The ground truth images are illustrated in Fig. 1(a). Based on the simulated data, the reconstructed images using the SENSE method are illustrated in Fig. 1(b).
Fig. 1(c) illustrates the reconstructed slices using our proposed method. It is worth noticing that 60 iterations were necessary to reach the convergence, which took us about 11 minutes with Matlab version $R2016a$ implementation (Processor Intel core i7- 7500U, up to 3.5GHz, RAM 8GB).
For the sake of comparison, results using a Tikhonov regularization [12] and a Bayesian ℓ_2 regularization [7] are provided. The reconstructed slices are displayed in Figs. 1(d) and 1(e), respectively.
From a visual point of view, we can easily notice that our sparse model gives a less noisy images compared to the other methods used for comparison. Moreover, it is clear that the used sparsity promoting prior (the BL model) helps retrieving images with low smoothing lever in comparison to the other methods.
As regards quantitative evaluation, Tab. I provides comparisons based on the signal to noise ratio and the structural similarity

criteria [13]. It is clear that the proposed method gives the less noisy images that are closer to the ground truth. These results are in a total agreement with the visual evaluation performed based on the reconstructed images.

TABLE I
SNR AND SSIM VALUES FOR THE TWO RECONSTRUCTED SLICES.

	Slice 1		Slice 2	
	SNR (dB)	SSIM	SNR (dB)	SSIM
SENSE	19.27	0.80	18.49	0.79
Prop. model	**28.85**	**0.95**	**27.15**	**0.94**
Tikhonov	21.31	0.90	20.48	0.89
Bay. ℓ_2 reg.	26.58	0.94	25.20	0.93

VI. CONCLUSION

In this paper, we present a new sparse Bayesian regularization technique for parallel MRI restoration using a Bernoulli-Laplace mixture prior that accounts for complex-valued data. The proposed model has been validated on a synthetic dataset. The obtained results show the good performance of our model for the processing of complex pMRI data. Our future work will focus on the validation of the proposed method on real data with more slices and acquisition configurations (number of coils, reduction factor,...).

REFERENCES

[1] D.J. Larkman and R.G. Nunes, "Parallel magnetic resonance imaging," *Physics in Medicine and Biology*, vol. 52, pp. 15–55, March 2007.

[2] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, pp. 952962, July 1999.

[3] M. Blaimer, F. Breuer, M. Mueller, R. M. Heidemann, M. A. Griswold, and P. M. Jakob, "SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method.," *Topics in magnetic resonance imaging : TMRI*, vol. 15, no. 4, pp. 223–236, august 2004.

[4] L. Chaari, J.-C Pesquet, A. Benazza-Benyahia, and P. Ciuciu, "Autocalibrated regularized parallel MRI reconstruction in the wavelet domain," in *IEEE International Symposium on Biomedical Imaging*, Paris, France, May 14-17 2008, pp. 756–759.

[5] L. Chaari, J.-C. Pesquet, A. Benazza-Benyahia, and P. Ciuciu, "A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging," *Medical Image Analysis*, vol. 15, pp. 185–201, November 2011.

[6] L. Chaari, P. Ciuciu, S. Mériaux, and J.-C Pesquet, "Spatio-temporal wavelet regularization for parallel MRI reconstruction: application to functional MRI," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 27, no. 6, pp. 509–529, December 2014.

[7] S. Chaabene and L. Chaari, "A Bayesian myopic parallel MRI reconstruction," *8th International Symposium on Signal, Image, Video and Communications (ISIVC)*, vol. 52, pp. 15–55, Tunis, Tunisia, November 21-23 2016.

[8] L. Chaari, J.-Y. Tourneret, and H. Batatia, "Sparse Bayesian regularization using Bernoulli-Laplacian priors," in *21st European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, September 9-13 2013, pp. 1–5.

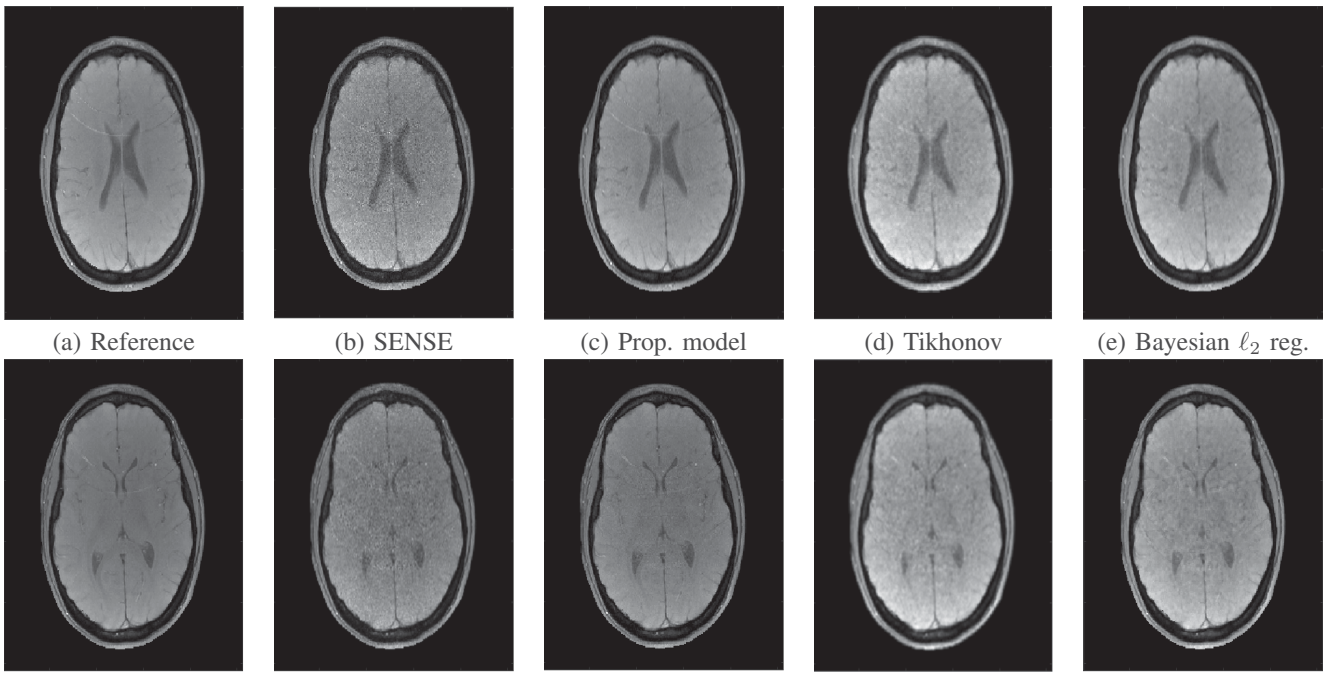
[9] L. Chaari, H. Batatia, and J.-Y. Tourneret, "Sparse Bayesian image restoration with linear operator uncertainties with application to EEG signal recovery," in *Middle East Conference on Biomedical Engineering (MECBME)*, Doha, Qatar, February 2014, pp. 139–142.

[10] F. Costa, H. Batatia, L. Chaari, and J.-Y. Tourneret, "Sparse EEG source localization using bernoulli laplacian priors," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 12, pp. 2888–2898, 2015.

[11] C. Robert and G. Castella, "Monte Carlo statistical methods," *Springer, New York*, 2004.

[12] A. Tikhonov, "Tikhonov Regularization of incorrectly posed problems," *Soviet Mathematics Doklady*, vol. 4, pp. 1624–1627, 1963.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.



(a) Reference (b) SENSE (c) Prop. model (d) Tikhonov (e) Bayesian ℓ_2 reg.

Fig. 1. Processed MRI slices: Reference slices (a), Reconstructed slices using SENSE (b), the proposed model (c), Tikhonov regularization (d), and the Bayesian ℓ_2 regularization algorithm (e).